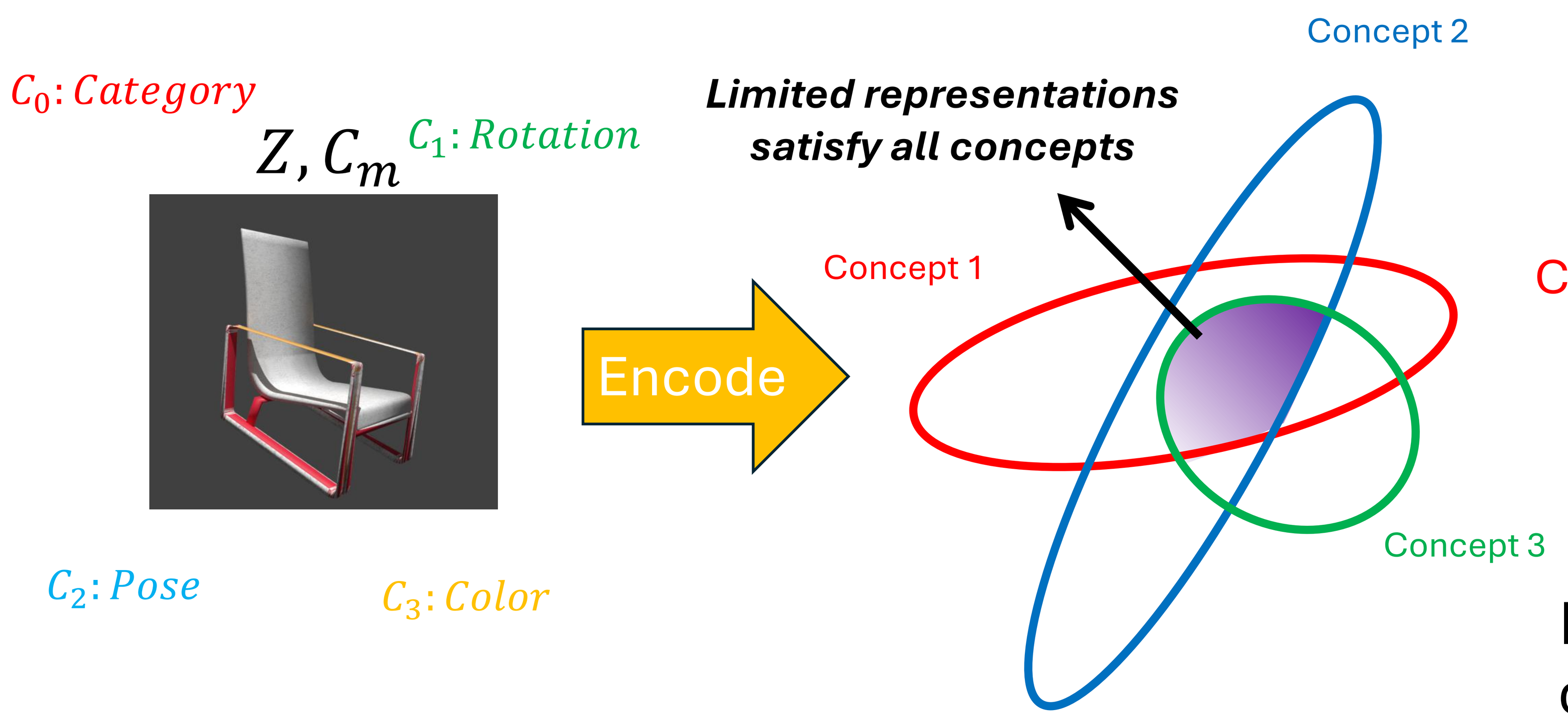
Domain Expansion: A Latent Space Construction Framework for Multi-Task Learning

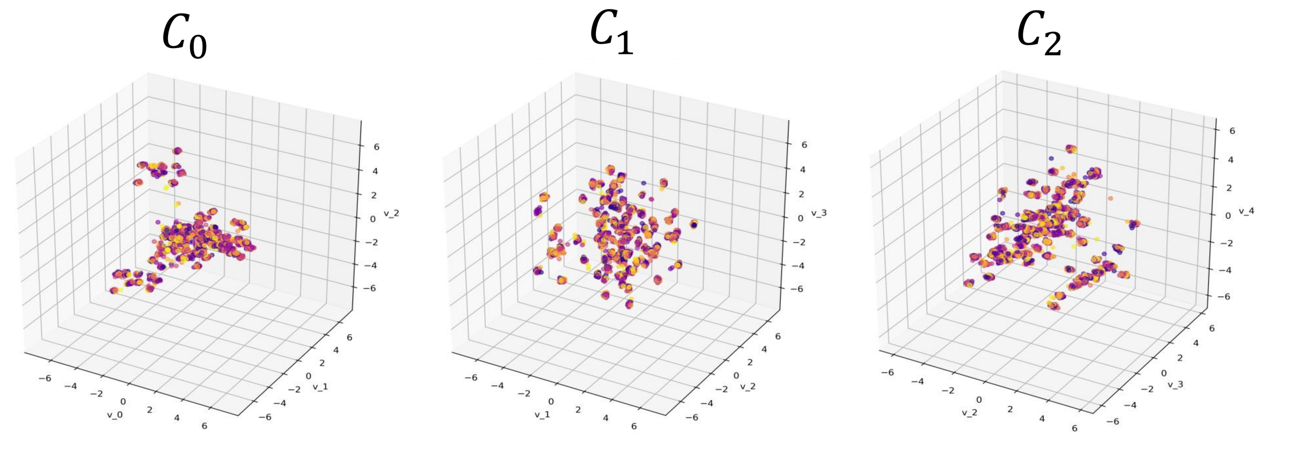
Chi-Yao Huang, Khoa Vo, Aayush Verma, Duo Lu, Yezhou Yang

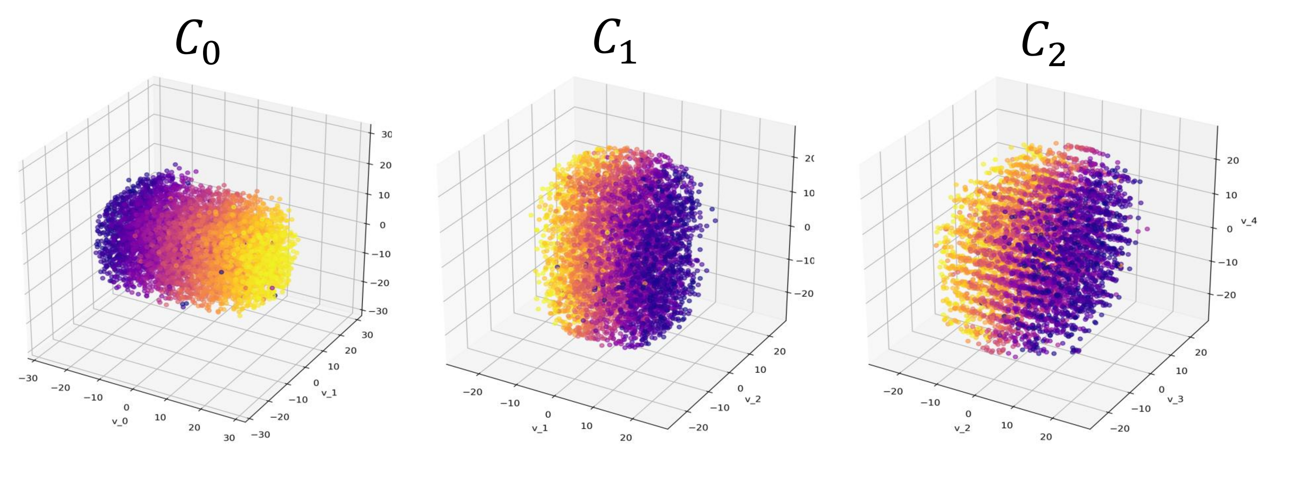
Arizona State University

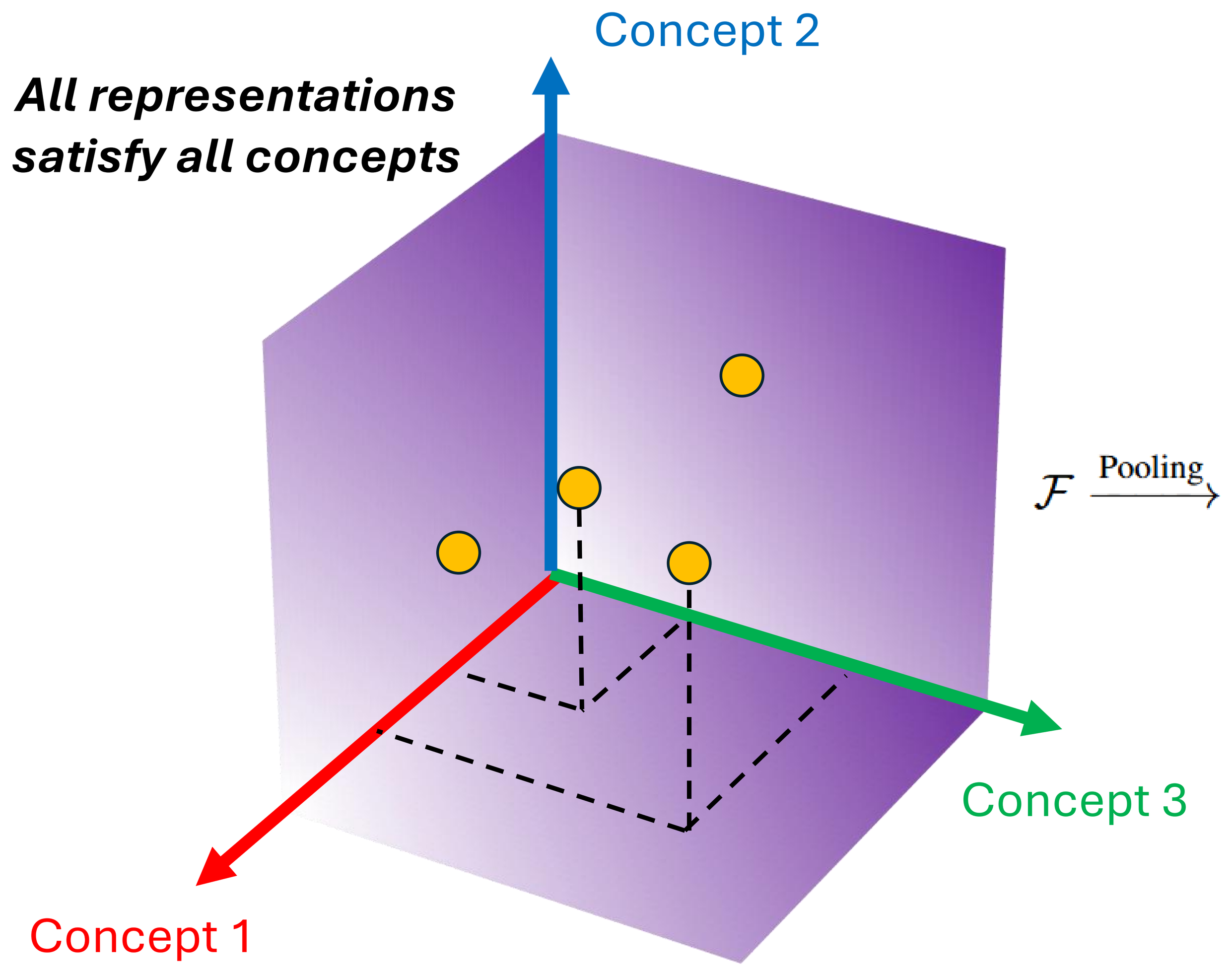
Latent Representation Collapse

A visual input often contains rich concepts. If we train a single model to handle all concepts at the same time, the features will be pulled in different directions, leading to a shrunken latent space. This shrunken latent space often degrades the model's performance. We call this phenomenon **latent representation collapse**.

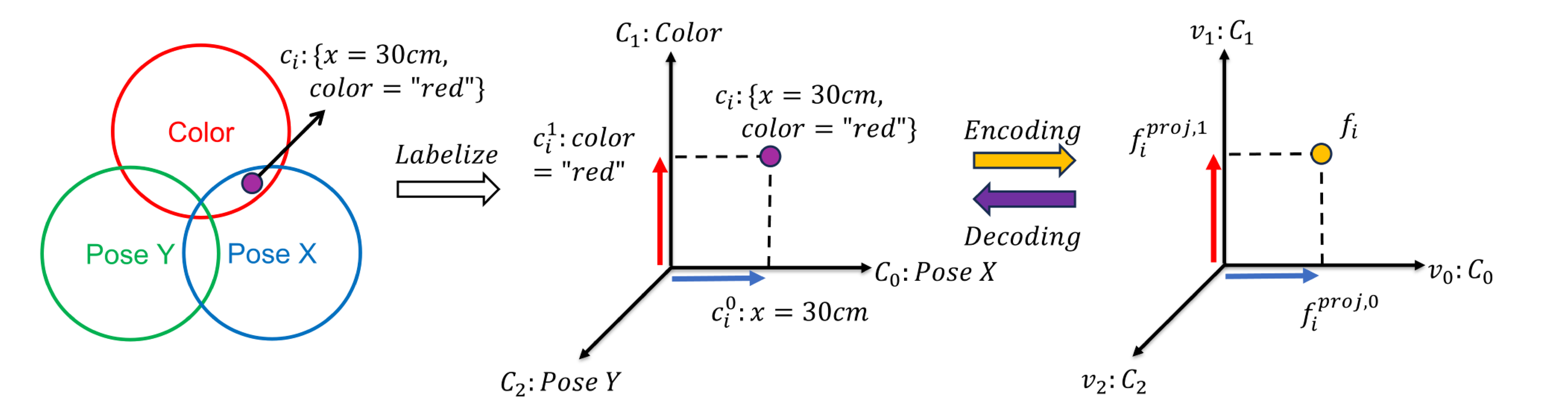






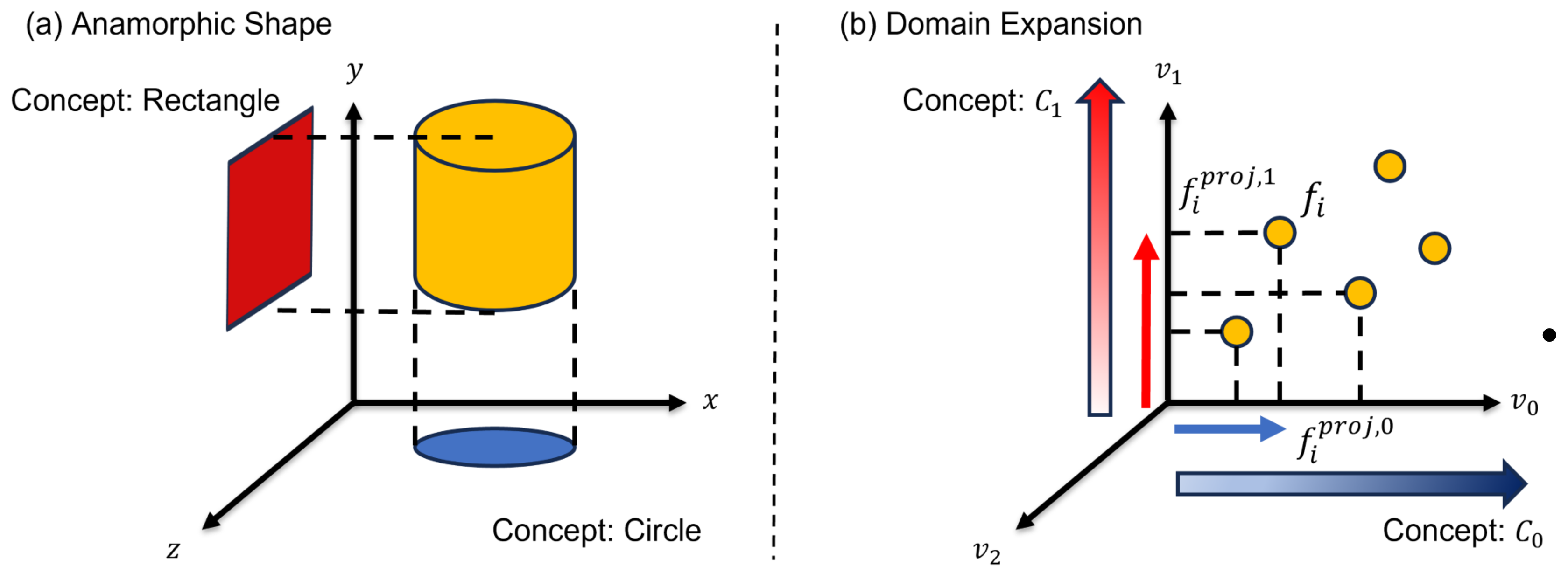


$$\mathcal{F} \xrightarrow{\text{Pooling}} \begin{cases} \mathcal{F}_0^{\text{proj}} \xrightarrow{\text{Dec}_0} \mathcal{C}_0 \\ \mathcal{F}_1^{\text{proj}} \xrightarrow{\text{Dec}_1} \mathcal{C}_1 \\ \vdots \\ \mathcal{F}_{M-1}^{\text{proj}} \xrightarrow{\text{Dec}_{M-1}} \mathcal{C}_{M-1} \end{cases}$$



Orthogonal Pooling

In the latent space, we use eigenvectors to represent concepts. Each eigenvector is assigned to a single target concept. This assignment ensures that the 1D subspace spanned by **the eigenvector is dedicated exclusively to each concept."**



Properties and Operators

- **Property 1: Multi-concept Encoding.**

$$f_i \xrightarrow{\text{Pooling}} \{f_i^{\text{proj},0}, \ldots, f_i^{\text{proj},M-1}\} \xrightarrow{\text{Dec}} \{c_i^0, \ldots, c_i^{M-1}\} \rightarrow c_i.$$

- **Property 2: Orthogonality of Target Concepts.**

$$\mathcal{F}_0^{\text{proj}} \perp \mathcal{F}_1^{\text{proj}} \perp \cdots \perp \mathcal{F}_{M-1}^{\text{proj}} \implies \mathcal{C}_0 \perp \mathcal{C}_1 \perp \cdots \perp \mathcal{C}_{M-1}.$$

- **Operator 1: Concept-Specific Adjustment ($\oplus^m$) and ($\ominus^m$).**

$$c_i \oplus^m c_\Delta^m \rightarrow \{c_i^0, \ldots, \{c_i^m \oplus^m c_\Delta^m\}, \ldots c_i^{M-1}\}$$

$$\xrightarrow{\text{Dec}^{-1}} \{f_i^{\text{proj},0}, \ldots, \{f_i^{\text{proj},m} + f_\Delta^{\text{proj},m}\}, \ldots, f_i^{\text{proj},M-1}\}$$

$$\xrightarrow{\text{Reconst}} f_i + f_\Delta^{\text{proj},m}.$$

- **Operator 2: Concept Composition (⊕) and (⊖).**

$$c_p \oplus c_q \rightarrow \{\{c_p^0 \oplus^0 c_q^0\}, \{c_p^1 \oplus^1 c_q^1\}, \ldots, \{c_p^{M-1} \oplus^{M-1} c_q^{M-1}\}\}$$

$$\xrightarrow{\text{Dec}^{-1}} \{\{f_p^{proj,0} + f_q^{proj,0}\}, \{f_p^{proj,1} + f_q^{proj,1}\}, \ldots, \{f_p^{proj,M-1} + f_q^{proj,M-1}\}\}$$

$$\xrightarrow{\text{Reconst}} f_p + f_q.$$

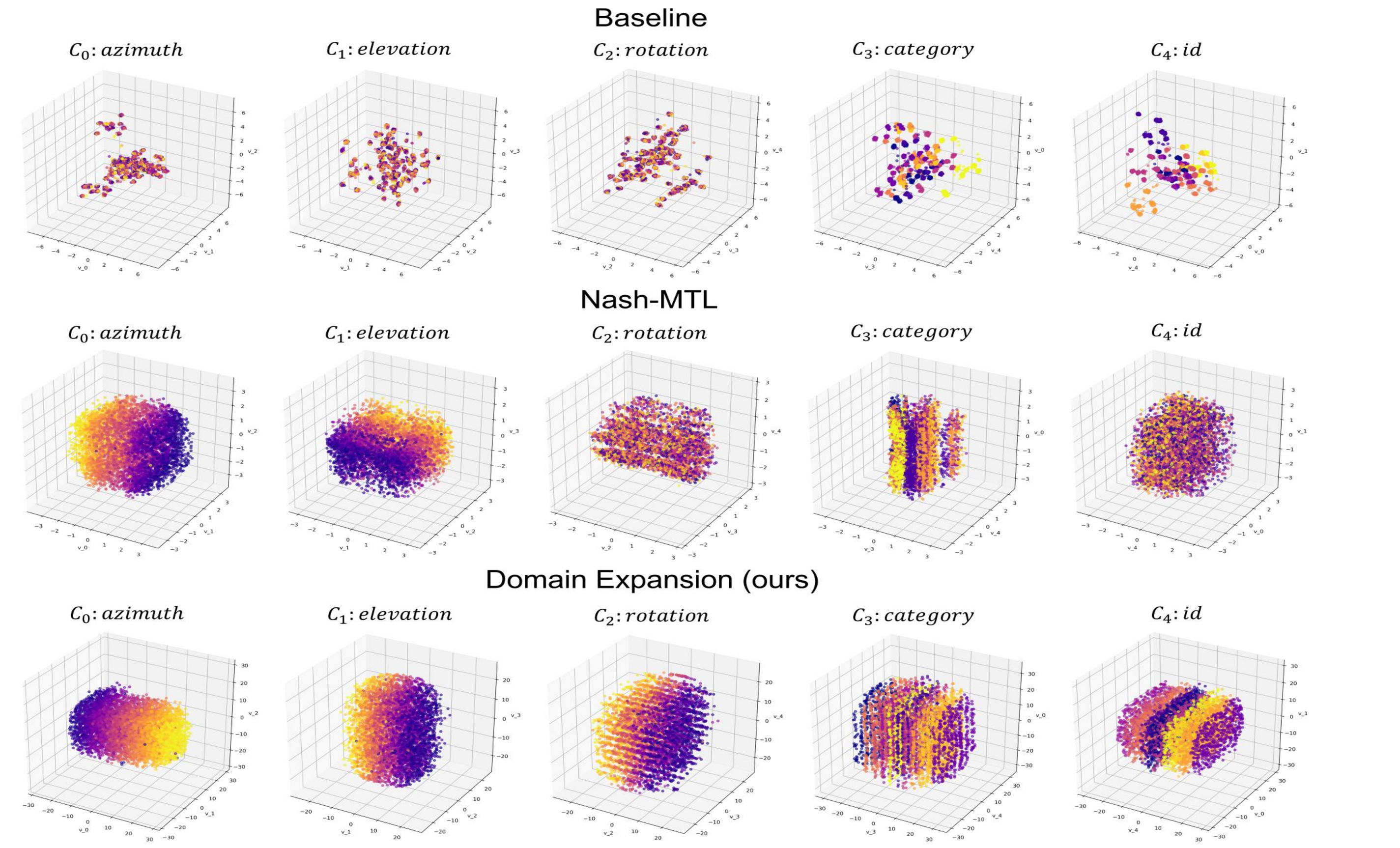


Table 1: Comprehensive comparison of representation quality, predictive performance, and concept composition. Arrows indicate whether higher (↑) or lower (↓) values are better.

Objective Set	Method	Representation & Predictive Performance										Concept Comp.
		Spearman ↑			V-score ↑		MAE° ↓			Acc. ↑		Sim. ↑
		az	el	rot	cat	id	az	el	rot	cat	id	⊕ and ⊖
Objective Set 1	baseline	0.41	0.34	0.35	0.16	0.14	0.12	0.09	0.09	0.28	0.37	0.22
	FAMO	0.49	0.41	0.42	0.00	0.00	0.12	0.09	0.09	0.19	0.18	0.28
	Nash-MTL	0.38	0.41	0.42	0.00	0.00	0.11	0.09	0.09	0.17	0.13	0.28
	IMTL	0.31	0.16	0.16	0.39	0.28	0.14	0.11	0.12	0.92	0.79	0.14
	Ours	0.95	0.87	0.85	0.99	0.91	0.08	0.08	0.09	0.99	0.97	0.95
Objective Set 2	baseline	0.01	0.01	0.01	0.99	0.00	0.77	0.38	0.38	0.99	0.99	0.42
	FAMO	0.28	0.23	0.22	0.99	0.00	0.19	0.14	0.13	0.99	0.99	0.28
	Nash-MTL	0.45	0.39	0.39	0.15	0.00	0.12	0.08	0.09	0.99	0.99	0.35
	IMTL	0.39	0.18	0.16	0.99	0.00	0.15	0.11	0.13	0.99	0.99	0.28
	Ours	0.95	0.87	0.85	0.98	0.96	0.07	0.08	0.09	0.98	0.94	0.93